

Eye-pupil displacement and prediction: effects on residual wavefront in adaptive optics retinal imaging

Caroline Kulcsár,^{1,*} Henri-François Raynaud,¹ and Aurea Garcia-Rissmann²

¹Laboratoire Charles Fabry – Institut d’Optique Graduate School – CNRS, Université Paris-Saclay, France

²Laboratório Nacional de Astrofísica /MCTI, Brazil – Previously with Université Paris 13, France

*caroline.kulcsar@institutoptique.fr

Abstract: This paper studies the effect of pupil displacements on the best achievable performance of retinal imaging adaptive optics (AO) systems, using 52 trajectories of horizontal and vertical displacements sampled at 80 Hz by a pupil tracker (PT) device on 13 different subjects. This effect is quantified in the form of minimal root mean square (rms) of the residual phase affecting image formation, as a function of the delay between PT measurement and wavefront correction. It is shown that simple dynamic models identified from data can be used to predict horizontal and vertical pupil displacements with greater accuracy (in terms of average rms) over short-term time horizons. The potential impact of these improvements on residual wavefront rms is investigated. These results allow to quantify the part of disturbances corrected by retinal imaging systems that are caused by relative displacements of an otherwise fixed or slowly-varying subject-dependent aberration. They also suggest that prediction has a limited impact on wavefront rms and that taking into account PT measurements in real time improves the performance of AO retinal imaging systems.

© 2016 Optical Society of America

OCIS codes: (110.1080) Active or adaptive optics; (170.4470) Ophthalmology; (120.0120) Instrumentation, measurement, and metrology; (070.2025) Discrete optical signal processing; (999.9999) Kalman filter; (999.9999) Eye pupil motion.

References and links

1. F. Roddier ed., *Adaptive Optics in Astronomy* (Cambridge University Press Cambridge, 1999).
2. J. Liand, B. Grimm, S. Goelz, and J.F. Bille, “Objective measurement of wave aberrations of the human eye with the use of a HartmannShack wave-front sensor,” *J. Opt. Soc. Am. A* **11**, 1949–1957 (1994).
3. J. Liang, D.R. Williams, and D.T. Miller, “Supernormal vision and high-resolution retinal imaging through adaptive optics,” *J. Opt. Soc. Am. A* **14**, 2884–2892 (1997).
4. A. Roorda, “Adaptive optics ophthalmoscopy,” *J. Refract. Surg.* **16**, S602–S607 (2000).
5. E.J. Fernández, I. Iglesias, and P. Artal, “Closed-loop adaptive optics in the human eye,” *Opt. Lett.* **26**(10), 746–748 (2001).
6. P. Godara, A.M. Dubis, A. Roorda, J.L. Duncan, and J. Carroll, “Adaptive optics retinal imaging: emerging clinical applications,” *Optometry Vision Sci.* **87**(12), 930–941 (2010).
7. S. Niu, J. Shen, C. Liang, Y. Zhang, and B. Li, “High-resolution retinal imaging with micro adaptive optics systems,” *Appl. Opt.* **50**(22), 4365–4375 (2011).
8. M. Ficocelli and F. Ben-Amara, “Online Tuning of Retinal Imaging Adaptive Optics Systems,” *IEEE Trans. Control Syst. Technol.* **20**(3), 747–754 (2012).

9. W. Zou, X. Qi and S. Burns, "Wavefront-aberration sorting and correction for a dual-deformable-mirror adaptive optics system," *Opt. Lett.* **33**(22), 2602–2604 (2008).
10. C. Li, N. Sredar, K.M. Ivers, H. Queener, and J. Porter, "A correction algorithm to simultaneously control dual deformable mirrors in a woofer-tweeter adaptive optics system," *Opt. Express* **18**(16), 16671–16684 (2010).
11. H. Hofer, P. Artal, B. Singer, J.L. Aragón, and D.R. Williams, "Dynamics of the eye's wave aberration," *J. Opt. Soc. Am. A Opt. Image Sci. Vis.* **18**(3), 497–506 (2001).
12. D.R. Iskander, M.J. Collins, M.R. Morelande, and M. Zhu, "Analyzing the dynamic wavefront aberrations in the human eye," *IEEE Trans. Biomed. Eng.* **51**(11), 1969–1980 (2004).
13. S. Gruppeta, F. Lacombe, and P. Puget, "Study of the dynamic aberrations of the human tear film," *Opt. Express* **13**(19), 7631–7636 (2005).
14. J.-F. Le Gargasson, M. Glanc, and P. Léna, "Retinal imaging with adaptive optics," *CR. Acad. Sci. Paris Applied Physics* **2**, 1131–1138 (2001).
15. S. Martinez-Conde, S. L. Macknik, and D. H. Hubel, "The role of fixational eye movements in visual perception," *Nat. Rev. Neurosci.* **5**, 229–240 (2004).
16. B. Sahin, B. Lamory, X. Levecq, F. Harms, and C. Dainty, "Adaptive optics with pupil tracking for high resolution retinal imaging," *Biomed. Opt. Express* **3**(2), 225–239 (2012).
17. D.X. Hammer, R.D. Ferguson, C.EE. Bigelow, N.V. Iftimia, T.E. Ustun, and S.A. Burns, "Adaptive optics scanning laser opthalmoscope for stabilized retinal imaging," *Opt. Express* **14**(8), 3354–3367 (2006).
18. J. Arines, P. Prado, and S. Bará, "Pupil tracking with a Hartmann-Shack wavefront sensor," *J. Biomed. Opt.* **15**(3), 036022 (2010).
19. A. Garcia-Rissmann, C. Kulcsár, H.-F. Raynaud, Y. El Mrabet, B. Sahin, and B. Lamory, "Adaptive prediction of human eye pupil position and effects on wavefront errors," *Proc. SPIE Ophthalmic Technologies XXI*, 78851W (2011).
20. L. A. Riggs, J.C. Armington, and F. Ratliff, "Motions of the retinal image during fixation," *J. Opt. Soc. Am.* **44**, 315–321 (1954).
21. Z. Ramdane-Cherif, A. Naït-Ali and O. Krebs, "An autoregressive (AR) model applied to eye tremor movement, clinical application in schizophrenia," *J. Med. Syst.* **28**(5), 489–495 (2004).
22. L.N. Thibos, X. Hong, A. Bradley, and X. Cheng, "Statistical variation of aberration structure and image quality in a normal population of healthy eyes," *J. Opt. Soc. Am. A* **19**, 2329–2348 (2002).
23. B. Sahin, B. Lamory, X. Levecq, L. Vabre, and C. Dainty, "Retinal imaging system with adaptive optics enhanced with pupil tracking," *Proc. SPIE* **7885**, 788517 (2011).
24. B. D. O. Anderson and J. B. Moore, *Optimal Filtering* (London Prentice Hall, 1979).
25. O. V. Komogortsev and J. I. Khan, "Eye movement prediction by oculomotor plant Kalman filter with brainstem control," *J. Cont. Th. Appl.* **7**(1), 14–22 (2009).
26. E. Walter and L. Pronzato, *Identification of Parametric Models from Experimental Data* (Springer UK, 1997).
27. Yi Cao, <http://www.mathworks.com/matlabcentral/fileexchange/18189-learning-the-extended-kalman-filter> (2009).
28. J.-R. Liang, S. Moshel, A. Z. Zivotofsky, A. Caspi, R. Engbert, R. Kliegl and S. Havlin, "Scaling of horizontal and vertical fixational eye movements," *Phys. Rev. E* **71**, 031909 (2005).
29. D. L. Sparks, "The brainstem control of saccadic eye movements," *Nat. Rev. Neurosci.* **3**, 952–964 (2002).
30. J. Jarosz, S. Meimon, J.-M. Conan, and M. Paques, "High temporal resolution ocular aberrometry with pupil tracking," *Proc. SPIE* 8930, *Ophthalmic Technologies XXIV*, 89300D (2014).
31. F. Felberer, M. Rechenmacher, R. Haindl, B. Baumann, C. K. Hitzenberger, and M. Pircher, "Imaging of retinal vasculature using adaptive optics SLO/OCT," *Biomed. Opt. Express* **6**(4), 1407–1418 (2015).
32. C. Viard, K. Nakashima, B. Lamory, M. Pques, X. Levecq, and N. Chteau, "Imaging microscopic structures in pathological retinas using a flood-illumination adaptive optics retinal camera," *Proc. SPIE* **7885**, 788509 (2011).

1. Introduction

Adaptive optics (AO) systems, which combine a wavefront sensor (WFS) and a deformable mirror (DM) inserted into the telescope's optical path, have been used since the early 1990s to counter the effects of atmospheric turbulence on ground-based telescopes [1]. This technique, together with associated DM and Hartmann-Shack (HS) WFS technology, has been subsequently and successfully adapted to correct optical aberrations in retinal imaging. In 1994, Liang et al demonstrated for the first time the feasibility of wavefront sensing in the eye [2]. This work was extended in 1997 by Liang and Williams, who measured the eye aberrations up to a very high order (65 Zernike modes) and were the first to close an AO loop on an eye in vivo and to obtain sharper images of a human retina [3] – see also [4, 5]. In the 2000s', as noted in a 2010 review of emerging clinical applications of this technique, 'AO imaging has

changed the way vision scientists and ophthalmologists see the retina, helping to clarify our understanding of retinal structure, function, and the etiology of various retinal pathologies' [6]. Existing retinal imaging AO systems use the same integral-action controller popular in astronomical AO to compute DM controls from WFS measurements. However, several more recent works have investigated and/or tested improved controller structures, including Smith predictors [7], adaptive controller tuning [8] and minimum-energy control for dual-deformable-mirror 'woofer-tweeter' systems [9, 10].

In astronomy, accurate models of atmospheric turbulence enable to construct detailed 'error budgets' which enable designers of new AO systems to translate image quality requirements of end-users (i.e., astronomers) into detailed performance requirements for all components and elements of the AO loop. Clearly, such an understanding of the physical nature and statistical properties of disturbances to be compensated would be hugely beneficial for future developments of retinal imaging AO systems. Studies of the trajectories of aberrations measured by HS WFSs suggested that different optical modes exhibit complex temporal behavior, hinting at a combination of diverse underlying mechanisms [11, 12]. Another experiment showed the non-negligible contribution of the tear film [13]. However, a number of authors noted that eye motion played a major role, see, e.g., [14, 15]. Thus, at least for the purpose of efficient DM control computation, a plausible conjecture is that a major part of performance degradation can be modeled as resulting from the relative displacements of a fixed or slow-moving and subject-dependent pupil aberration, as stated in [16]. One way to investigate the implications of this conjecture in terms of achievable performance is to consider an ideal case where the aberration seen by the AO system results only from the relative horizontal and vertical relative displacements (with respect to the imaging system) of an otherwise fixed pupil aberration. Throughout this paper, we shall call this simplified scenario the 'moving aberration' assumption.

For an ideal retinal imaging AO controller achieving perfect sensing of the aberration through the WFS and perfect compensation by the DM, the residual variance would be equal to the tracking error variance. In addition, under the 'moving aberration' assumption, this tracking error variance can be evaluated by taking a representative sample of fixed aberrations and moving them by appropriately distributed horizontal and vertical displacements during the total AO loop's delay between WFS measurement and DM correction.

The moving aberration hypothesis also suggests that a sensor capable of monitoring pupil displacements in real time could be used to improve aberration correction. In recent years, such so-called 'pupil trackers' (PT) camera-based devices have been integrated into retinal imaging systems and have been used for a number of purposes, for example to detect and forecast where a subject is looking within a scene. In 2006, Hammer et al. used a pupil tracker integrated into a scanning laser ophthalmoscope system to actively compensate pupil movements in real time using a flat two-degrees of freedom field stabilization mirror [17]. A WFS-based pupil tracker, which eliminates the need for a separate PT camera, was proposed in [18].

In 2012, Sahin et al. [16] implemented a PT-based real-time control scheme where DM's inputs were computed by shifting a previously estimated eye aberration across the horizontal and vertical axis according to pupil displacement measurements. This experiment conducted on a robotic 'model eye' and three human test subjects showed that the PT-based control achieved a level of correction performance broadly similar to a conventional AO controller. This pioneering work provided additional (if somewhat circumstantial) evidence in support of the moving aberration assumption. This PT-based control experiment used a PT system developed by the company Imagine Eyes.

The analysis presented in this paper extends the preliminary results obtained in [19], with more statistical considerations and in-depth performance evaluation. We used a set of pupil-displacement trajectories measured by the same PT device (see Acknowledgments section),

together with measured wavefronts, to derive quantitative assessments of the tracking error variance term entering an AO error budget under the moving aberration assumption. The displacement data set comprises a total of 52 displacement trajectories, each about 13 seconds long, recorded by Betül Sahin at Imagine Eyes on 13 different healthy subjects, with a PT based on a camera running at 80 frames per second. The wavefront data set corresponds to 500 wavefronts measured on healthy subjects by Imagine Eyes, under the form of Zernike coefficients for radial orders 1 to 6. Using these data sets, we simulate and quantify the tracking error due to the displacement of the aberration between the time its last position has been measured by the system and the moment the DM correction is applied. As explained above, this corresponds to the ultimate performance (expressed in terms of residual wavefront rms) that could be achieved by a standard AO loop fitted with ideally accurate WFS and DM and operating at a given WFS/DM sampling rate – or alternatively by a purely PT-based controller endowed with an exact knowledge of the fixed aberration.

In both cases, the tracking error is expressed centrally as a function of the total delay between the last available WFS/PT acquisition and DM correction – in our case, for delays ranging from 12.5 ms to 85 ms. Also, this computation is predicated on the implicit assumption that the last available pupil position measurement is used to estimate the pupil position at the time when the control is actually applied. An alternative would be to use a more elaborate method to predict the pupil's displacement in this short time interval. To this effect, we have tested several standard short-time prediction methods based on simple dynamic stochastic models of eye pupil displacement.

In the domain of retinal imaging, whenever a patient is fixing on a target, involuntary residual movements are indeed still present, due to a combination of eye and head movements. For instance, the human eye is known to have micro-movements [15] which are essential to our vision as they make it possible to maintain the transmission of the image fixed in our brain [20]. Such micro-movements are for example studied in [21], where the authors seek a mono-dimensional Auto-Regressive (AR) model of tremor movements. Martinez-Conde et al. [15] mention three main types of eye movements during visual fixation: tremor, drifts and micro-saccades. Quantitatively, pupil displacements occur with different translational/torsional amplitudes and frequencies (comprising approximately 0.001-1 degrees and 0.1-100 Hz), and are known to vary significantly from individual to individual [22, 18], and also for one same individual observed at different times. Dynamic stochastic models of eye pupil displacement need therefore to be adapted to each individual, and thus identified from each displacement trajectory.

This paper is organized as follows. Section 2 describes the PT set-up and the data sets used throughout the paper; a preliminary statistical analysis of the pupil displacement trajectories is presented, focusing on the issues of saccade detection in real time and PT frame-rate selection. Section 3 is devoted to construction and identification from PT data of stochastic pupil displacement models. Section 4 discusses the performance of these identified models for short-term prediction of pupil displacements. In section 5, the tracking errors (residual wavefront rms) corresponding to the recorded pupil displacement trajectories applied to the measured wavefront data set are evaluated, and the impact of pupil prediction on performance is discussed. Some conclusions and perspectives for future works are presented in Section 6.

2. Description of the PT data and basic setup

The pupil tracking system (see [23] for more details) consists in imaging the pupil with a CCD camera using Near Infrared (NIR) LEDs as light sources, followed by digital image post-processing. The system works for 6-9mm diameter pupils with an accuracy of $\pm 20 \mu\text{m}$ (3σ). The measurement sampling frequency F_s is 79 Hz or 80 Hz (depending on the acquisitions), with a CCD exposure time of 10 ms. The subject's head is stabilized with a standard ophthalmic

chinrest (see Fig. 1). Measurements from this PT correspond to eye pupil's positions following the horizontal and vertical axes. Our PT dataset consists of 52 trajectories, corresponding to data acquired from 13 different people, each trajectory being about 13s long.

Figure 2 shows an example of position measurements for a subject in our sample. These horizontal and vertical trajectories show abrupt changes in measured positions due to saccades (zones delimited by the dashed lines), which can also occur after eye blinking. These data are used for a study on short-term prediction of eye pupil positions, where the identified dynamic stochastic models allow also for automatic detection of saccades in real time. Then, we evaluate the impact on residual wavefronts of pupil displacement and its predictions. A total of 500 phase screens measured on different subjects by Imagine Eyes using a 32×32 high-resolution Shack-Hartmann (haso 32-eye Wavefront Sensor) were used for the study. They have been compared with synthetic ones obtained by Thibos' model [22] in Section 5.

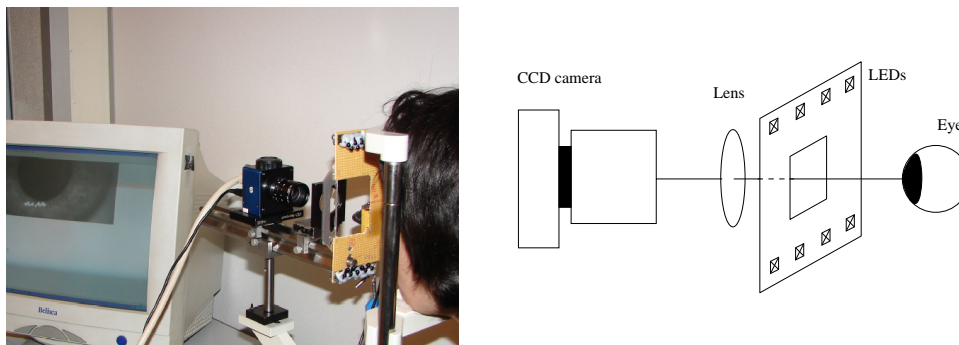


Fig. 1. Pupil-Tracker (PT): acquisition (left) and system's description.

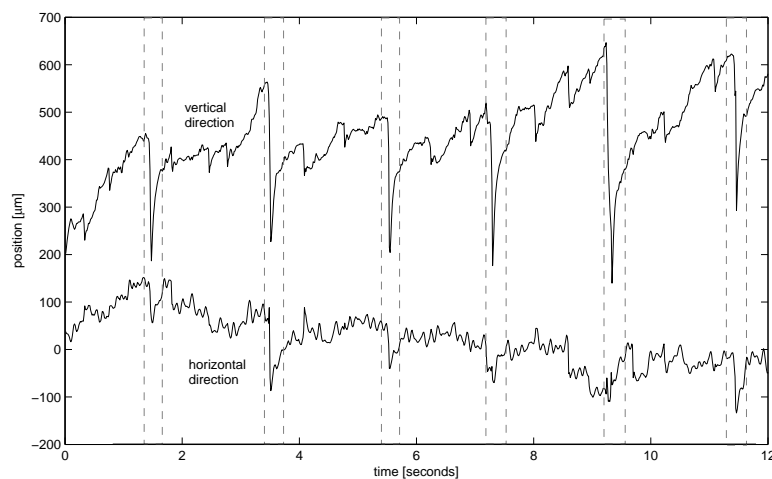


Fig. 2. Example of pupil position measurements for one of the subjects in our sample (dashed lines delimit saccades zones).

We present in Fig. 3 two trajectories of horizontal and vertical pupil displacements, corresponding to the same subject, during fixation of a fixed target. This illustrates the high variabil-

ity of behavior for a same subject, and the presence of different types of eye micro-movements noted in the introduction. These can be described as follows:

- **Drifts:** slow motions of the eye (up to a few Hertz), average 2-5 arcmin in amplitude with a mean speed of about 6 arcmin/s.
- **Tremors:** very high frequency oscillations. Tremors' amplitudes and frequencies are usually in the range of the recording system's noise. Motions amplitudes average 5-60 arcsec at frequencies up to ≈ 90 Hz.
- **Micro-saccades:** jerk-like eye movements averaging 6 arcmin in amplitude.

In these two recordings, one can distinguish clearly between blinks and saccades. A blink of the eye is separated from a saccade by the absence of measurements from the PT. Tremors correspond to smaller amplitude movements. It may be also noted that after a blink the position may not be similar to that before the blink. In other words, a blink goes often with a saccade. Also, a saccade following the horizontal axis does not necessarily correspond to a saccade following the vertical axis, and *vice versa*.

With the selected sampling frequency of about 80 Hz, there will be aliasing on the tremor's and micro-saccades' spectra. (Shannon-Nyquist sampling theorem: the sampling frequency F_s must be greater than twice the maximum frequency of the analog signal to avoid aliasing of frequencies $> F_s/2$ and to allow a perfect reconstruction from the sampled values.) In the case of tremor, its very low amplitude (less than the diameter of a cone, as reported in [15]) leads to include a large part of this movement in the system's noise range (20 μm), so that the aliasing of frequencies above 40 Hz will have anyway a low effect. Micro-saccades have a random occurrence and a too short duration to be corrected in such an AO loop, but they are detectable thanks to their high amplitudes and will be regarded as outliers values.

It has to be noted that measurements made by the PT result from the computation of the center of the pupil in the PT plane, the pupil positions resulting from the projection on this plane of a combination of eye and head movements. The center of the pupil is calculated taking the center of an ellipse fitted to the image of the pupil taken by the PT's camera [23]. Head movements may also induce in the measurements slow drifts and kinds of saccades, which become indistinguishable from the movements produced by the sole eye.

According to Thibos [22], one subject can present a high diversity in pupil movements, therefore we consider all the PT data analyzed here as independent from each other. Moreover, instead of modeling the eye pupil's positions, we will consider one-step displacements (i.e. differences between positions). This eliminates the need to estimate the absolute reference position, enables prediction algorithms to lock in immediately when saccades occur, and allows to model the process around a zero value. Moreover, in an AO system, only displacements are necessary to update the control values. As these displacements are computed from noisy position measurements, they will be in return affected by a stronger noise component.

Prior characterization of the movements amplitude is necessary for the improvement of the position estimation process. Let us here consider all one-step displacements as an ensemble. Displacements depend on (and increases with) the time increment for which they are evaluated. We can then plot the absolute value of the one-step displacements as a function of the time delay between PT measurements. Figure 4 shows the upper bounds within which lie 99%, 95% and 90% of the absolute value of one-step displacements (dashed, dot-dashed and dotted lines, respectively), and also the curve of their average values (dot-dashed line), as a function of the delay between measurements; the plots mean that the amplitudes of practically all one-step displacements are concentrated below 200 μm , and that extreme changes in the trajectories represent a few events. Also, it can be seen that at $\Delta t = 12.5$ ms, 95% of the one-step displace-

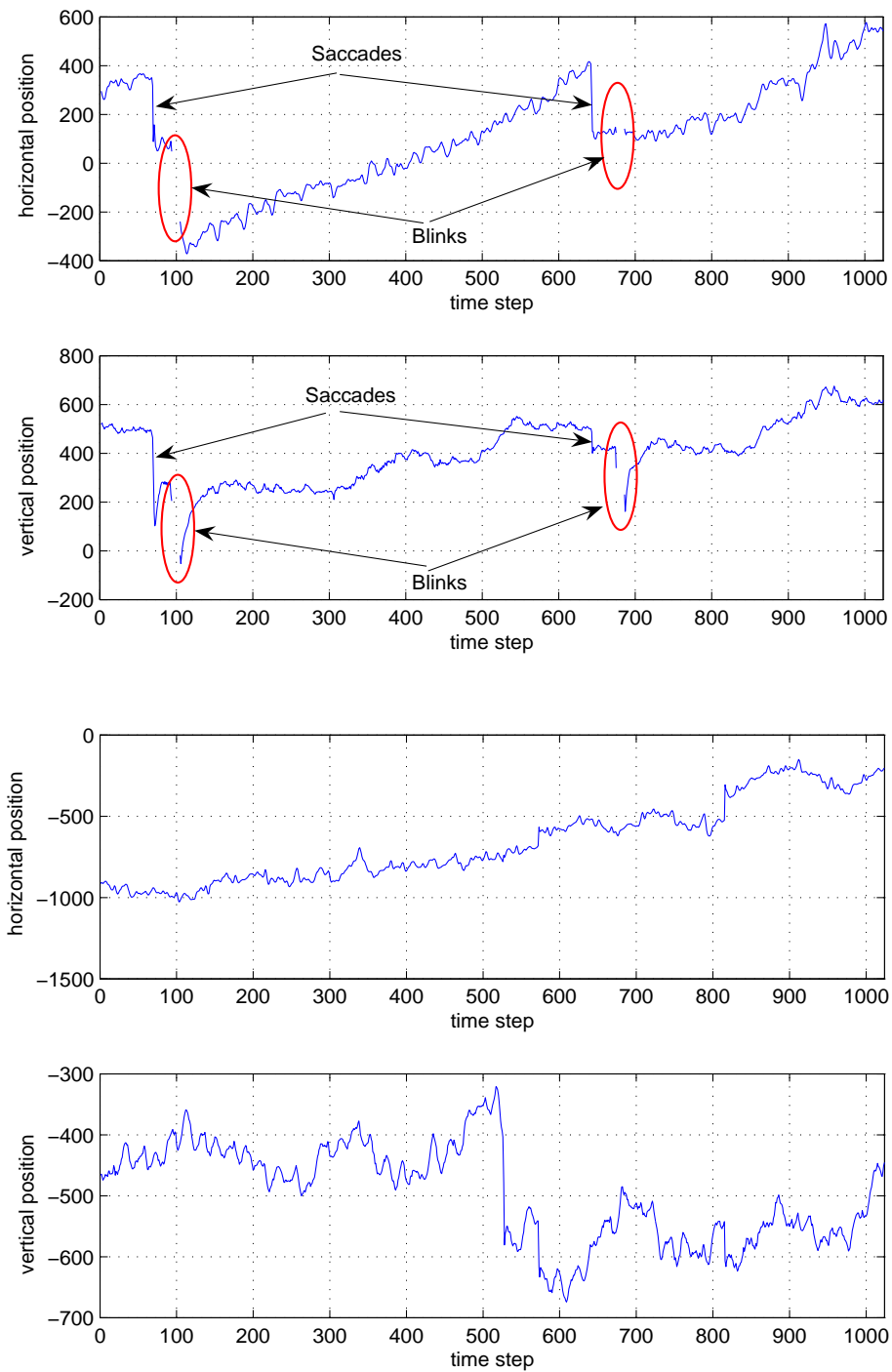


Fig. 3. Example 1: PT measurements during fixation for the same subject at different instants (top and bottom). The time unit is $\Delta t = 12.5$ ms. Signal is a mixture of saccades, drift, tremor, blink and small head movements. Noise measurement (3σ) is $20\ \mu\text{m}$.

ments have a value in the range of the PT measurement noise (20 μm). Thus, with this noise level, it would be pointless to increase the PT rate.

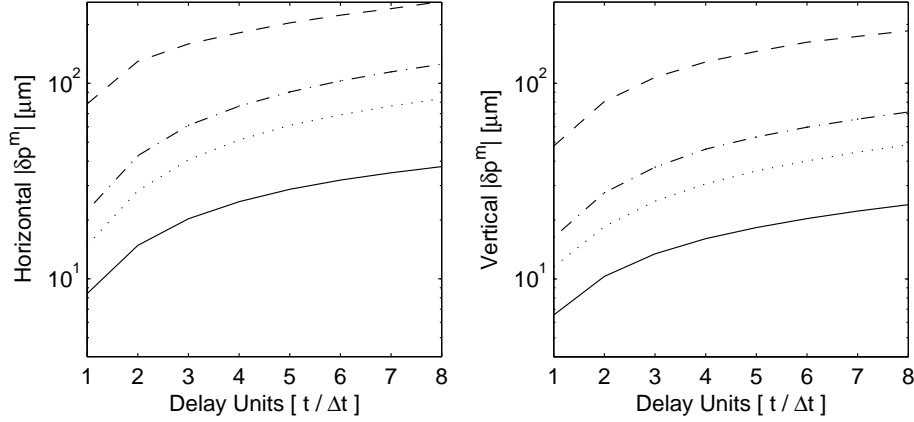


Fig. 4. Statistics of displacements as a function of the delay between PT measurements. The unit delay is $\Delta t = 12.5$ ms. The curves represent the displacement values δp^m under which we have 99%, 95% and 90% (dashed, dot-dashed and dotted lines, respectively) of the data, when the delay between two successive positions varies from Δt to $8\Delta t$. The average value calculated considering all data is also shown (solid line).

3. Models and algorithms for displacement prediction

The choice of a model implies a choice of the most suitable variables to work on. As noted above, we will consider *one-step displacements* defined as

$$\delta p_k = p_k - p_{k-1} \quad (1)$$

where p_k is the vector of horizontal and vertical positions at sample time $k\Delta t$. Likewise, the measured one-step displacement δp_k^m is defined as:

$$\delta p_k^m = p_k^m - p_{k-1}^m \quad (2)$$

where p_k^m is the PT measurement at time $k\Delta t$. In order to correct for aberrations, one needs to predict the actual position p . However, in order to assess the quality of this prediction, one would need to have access, at least afterwards, to the true values of the eye pupil positions. In our setup, this ground truth is not available to evaluate prediction performance. Performance is thus evaluated by comparing predictions with future measurements. Assuming that the measurement noise is an additive zero mean white noise with standard deviation σ for both horizontal and vertical positions, the performance indicators based on position errors are affected by an additional uncertainty. This will not have a significant impact on relative performance of prediction methods: the squared rms of the prediction error is equal to $rms^2 = \|\hat{p} - p^*\|^2 + \sigma^2$ where p^* is the true position (which is unknown). Therefore, the measurement noise affects the rms computation in the same way for all methods.

As a consequence of Eq. (2), the position at time index $k + \ell$, $\ell \geq 1$, is given by

$$p_{k+\ell} = p_k + \sum_{j=1}^{\ell} \delta p_{k+j}. \quad (3)$$

An estimate of $p_{k+\ell}$ based on all the measurements available until time k will then be computed as

$$\hat{p}_{k+\ell|k} = p_k^m + \sum_{j=1}^{\ell} \widehat{\delta p}_{k+j|k} \quad (4)$$

where $\widehat{\delta p}_{k+j|k}$ is the predicted values of the one-step displacements.

We will consider in the following three different algorithms for predicting pupil displacements: the reference one, called ‘dummy’, consists in doing nothing else than taking as a prediction the last available measurement. This is the reference because it is the simplest, and because when the PT frame is of the order of an AO system frequency loop, the dummy predictor is equivalent to the ideal AO loop, where perfect WFS measurements are used for perfect DM correction. We will also consider two predictors based on Kalman filters, with simple autoregressive models of order 1.

3.1. A simple (dummy) predictor

Since one-step displacements are weakly correlated, a simple way to predict the future positions is to take $\widehat{\delta p}_{k+j|k} = 0$ for any $j \geq 1$. This leads to take the last measured position p_k^m as predicted value for all future positions:

$$\hat{p}_{k+\ell|k} = p_k^m. \quad (5)$$

The underlying position model is a random walk driven by a white noise sequence. This will hereafter be called the dummy predictor.

3.2. Observer with parameter estimation

We reformulate the prediction problem in standard state-space form:

$$x_{k+1} = A x_k + v_k \quad (6)$$

$$y_k = C x_k + w_k \quad (7)$$

where v and w are independent Gaussian white noises with covariance matrices Σ_v and Σ_w respectively. The vector y_k denotes the measured output, in our case $y_k = \delta p_k^m$. Assuming uncorrelated horizontal and vertical one-step displacements measurements with same variance, we take here $\Sigma_w = \sigma_w^2 I$. The so-called state vector x contains the quantities that need to be estimated. In order for x and y to be stationary processes, A should have all its eigenvalues inside the unit circle. The magnitude of the eigenvalues expresses how fast the variable x decorrelates in time. A simple autoregressive model of order 1 (AR1) has been chosen for both horizontal and vertical one-step displacements. A scalar AR1 process $\{\eta\}$ is defined as

$$\eta_{k+1} = a\eta_k + v_k \quad (8)$$

where v is a zero mean Gaussian white noise and $|a| < 1$. Using such a model for horizontal and vertical displacements leads to $x = \delta p$ and

$$A = \begin{pmatrix} a_h & 0 \\ 0 & a_v \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \quad (9)$$

Assuming uncorrelated horizontal and vertical displacements, covariance matrix Σ_v is taken as $\Sigma_v = \sigma_v^2 I$.

When A and C are known, the state x can be estimated optimally (in the sense of minimizing the variance of the estimation error) using an optimal observer, the widely used Kalman filter

(see, e.g., [24], or [25] for an example in the different context of gaze prediction and anatomical-based models). It shall thereafter be referred to as the SKF, for *standard* Kalman filter. The SKF is described by recursive equations summarized as

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + H_k (y_k - C \hat{x}_{k|k-1}) \quad (10)$$

$$\hat{x}_{k+1|k} = A \hat{x}_{k|k} \quad (11)$$

where the so-called Kalman gain H_k is itself computed recursively together with $\Sigma_{k+1|k}$, the covariance matrix of the prediction error (see below).

As said above, this formulation requires the matrix A to be known. This is not the case here and therefore an on-the-fly identification is required along with the prediction. In our case, matrix A contains two unknown parameters, a_h and a_v . The parameter vector to be estimated is therefore:

$$\theta = \begin{pmatrix} a_h \\ a_v \end{pmatrix} \quad (12)$$

and we shall denote as $A(\theta) = \text{diag}(\theta)$ the corresponding value of A . Two solutions with low computational burden have been tested: Kalman filtering with separate parameter identification, and extended Kalman filter.

3.2.1. Standard Kalman filter with recursive least squares

We now use a recursive least-squares (RLS) procedure (see, e.g., [26]) to update the estimation of parameter θ based on all the data obtained until time $k\Delta t$. This method is applicable to models that are linear in parameters, *i.e* with measurements that can be expressed as

$$y_k = \delta p_k = r_k \theta + w_k. \quad (13)$$

For our AR1 models with independent displacements, r_k is a diagonal matrix with diagonal terms δp_{k-1} , and the covariance matrix of $\{w\}$ is also diagonal. Thus, the RLS algorithm can be split into two independent scalar estimators. We give here the more general version which gives at each time step k an estimated value $\hat{\theta}_k$:

1. Get y_k (new measurement).
2. Using the previous estimate $\hat{\theta}_{k-1}$, calculate the output error $e_k = y_k - r_k \hat{\theta}_{k-1}$.
3. Compute the estimation error covariance matrix Σ_k^{mc} :

$$\Sigma_k^{\text{mc}} = \Sigma_{k-1}^{\text{mc}} - \frac{\Sigma_{k-1}^{\text{mc}} r_k r_k \Sigma_{k-1}^{\text{mc}}}{1 + r_k \Sigma_{k-1}^{\text{mc}} r_k} \quad (14)$$

4. Define the gain L_k as $L_k = \Sigma_k^{\text{mc}} r_k$.
 5. Compute the new estimate as
- $$\hat{\theta}_k = \hat{\theta}_{k-1} + L_k e_k$$
6. Increment k and go to step 1.

The algorithm starts at $k = k_0$ such that the matrix $r_{k_0} r_{k_0}$ is invertible, and with initial values:

$$\Sigma_0^{\text{mc}} = (r_{k_0} r_{k_0})^{-1}$$

$$\hat{\theta}_0 = \Sigma_0^{\text{mc}} r_{k_0} y_{k_0}$$

with

$$\underline{r}_{k_0} = \begin{pmatrix} r_0 \\ \vdots \\ r_{k_0} \end{pmatrix}, \underline{y}_{k_0} = \begin{pmatrix} y_0 \\ \vdots \\ y_{k_0} \end{pmatrix}. \quad (15)$$

One iteration of the SKF+RLS identification/prediction procedure can then be summarized as follows:

1. Get y_k (new measurement).
2. Using the previous estimate $\hat{x}_{k|k-1}$, calculate the innovation $\tilde{y}_k = y_k - C \hat{x}_{k|k-1}$.
3. Compute the estimation gain $H_k = \Sigma_k C^T (\Sigma_w + C \Sigma_k C^T)^{-1}$.
4. Update the state estimate $\hat{x}_{k|k} = \hat{x}_{k|k-1} + H_k \tilde{y}_k$.
5. Update $\Sigma_{k|k} = \Sigma_k - H_k C \Sigma_k$.
6. Run a RLS to update parameter $\hat{\theta}_k$.
7. Calculate $\Sigma_{k+1} \equiv \Sigma_{k+1|k} = A(\hat{\theta}_k) \Sigma_{k|k} A(\hat{\theta}_k)^T + \Sigma_v$.
8. Compute one-step and two-step ahead predictions $\hat{x}_{k+1|k} = A(\hat{\theta}_k) \hat{x}_{k|k}$, $\hat{x}_{k+2|k} = A(\hat{\theta}_k)^2 \hat{x}_{k|k}$.
9. Increment k and go to step 1.

Initialization of all variables is set to zero, except for covariance matrices which are set to λI , with $\lambda \gg 1$. Note that $\hat{x}_{k+1|k}$ and $\Sigma_{k+1|k+1}$ are no longer optimal nor do they correspond anymore to conditional expectations; $\hat{x}_{k+1|k}$ is only an estimate based on all measurements until time index k . This will be true also for the extended Kalman filter described below.

When using the SKF+RLS method, we have set some limits to the estimated transition matrix diagonal values. (This matrix is 2×2 , comprising the two axes of movement.) They were forced to lie between 0.3 and 0.9 (a value of one is equivalent to the random walk case), and then the transition state matrix was updated (at each time step) taking the average of the 40 latest estimated values for its diagonal elements. This ensured that the parameter values would not diverge.

Also, in order to get the most out of the models described in Section 3 we had to scan through a set of input parameters, such as the relative strength of the noise and process covariances matrices. We screened through a wide range of values, and found that by taking $\sigma_v^2 / \sigma_w^2 \approx 100$ we obtained the best results in terms of rms of the difference between prediction and measured data. Performance is however not very sensitive to this value. It should nevertheless not be too small, otherwise filters may be destabilized.

3.2.2. Extended Kalman filter

Another classical approach, very close to the previous one, consists in using an extended Kalman filter (EKF, see e.g. [24]). In this formulation, parameters to be estimated are also considered as state variables, so that an extended state denoted by x^e is formed:

$$x_k^e = \begin{pmatrix} x_{1,k}^e \\ x_{2,k}^e \end{pmatrix} = \begin{pmatrix} x_k \\ \theta \end{pmatrix} \quad (16)$$

Based on Eq. (6)-(7), we can write:

$$x_{1,k+1}^e = A(x_{2,k}^e)x_{1,k}^e + v_k \quad (17)$$

$$x_{2,k+1}^e = x_{2,k}^e \quad (18)$$

$$(19)$$

which can be rewritten in a matrix compact form as

$$x_{k+1}^e = f^e(x_k^e) + v_k^e \quad (20)$$

where

$$f(x_k^e) = \begin{pmatrix} A(x_{2,k}^e)x_{1,k}^e \\ x_{2,k}^e \end{pmatrix} \quad \text{and} \quad v_k^e = \begin{pmatrix} v_k \\ 0 \end{pmatrix} \quad (21)$$

The function $f^e(x_k^e)$ mixes the data and parameters in a nonlinear way. The extended Kalman filter based on such a system corresponds to an observer in the form:

$$\hat{x}_{k|k}^e = \hat{x}_{k|k-1}^e + L_k \left(y_k - f^e(\hat{x}_{k|k-1}^e) \right) \quad (22)$$

$$\hat{x}_{k+1|k}^e = f^e(\hat{x}_{k|k}^e) \quad (23)$$

The observer gain L_k is computed using a first order Taylor expansion of f^e around the current estimated value:

$$f^e(x_k^e) \approx f^e(\hat{x}_{k|k}^e) + \mathcal{A}_k (x_k^e - \hat{x}_{k|k}^e) \quad (24)$$

where \mathcal{A}_k is the Jacobian matrix evaluated in $\hat{x}_{k|k}^e$ and defined as

$$\mathcal{A}_k = \left. \frac{\partial f^e(x_k^e)}{\partial x_k^e} \right|_{\hat{x}_{k|k}^e} \quad (25)$$

In the case of the AR1 models in Eq. (6), we have

$$\mathcal{A}_k = \begin{pmatrix} A(\hat{x}_{2,k|k}^e) & \text{diag}(\hat{x}_{1,k|k}^e) \\ 0 & I \end{pmatrix} \quad (26)$$

The gain matrix L_k is computed using this linear approximation and the measurement Eq. (7). Using the extended state, this equation can be rewritten as

$$y_k = \mathcal{C} x_k^e + w_k \quad (27)$$

with $\mathcal{C} = (I \ 0)$. One iteration of the complete procedure can be summarized as:

- Given the previously estimated state vector $x_{k|k-1}^e$ and a new measurement, calculate the innovation $\tilde{y}_k = y_k - \mathcal{C} \hat{x}_{k|k-1}^e$.
- Update the state $\hat{x}_{k|k}^e = \hat{x}_{k|k-1}^e + H_k \left(y_k - \mathcal{C} \hat{x}_{k|k-1}^e \right)$.
- Compute \mathcal{A}_k using Eq. (26).
- Calculate $\Sigma_{k+1} \equiv \Sigma_{k+1|k} = \mathcal{A}_k \Sigma_k \mathcal{A}_k^T + \Sigma_v$.
- Update the gain: $H_{k+1} = \Sigma_{k+1} \mathcal{C}^T (\Sigma_w + \mathcal{C} \Sigma_{k+1} \mathcal{C}^T)^{-1}$.
- Calculate $\Sigma_{k+1|k+1} = \Sigma_{k+1|k} - H_{k+1} \mathcal{C} \Sigma_{k+1|k}$.

The updated state vector and covariance matrix of the prediction error is calculated through an algorithm adapted from Yi Cao [27]. It is worth noticing that an extension to any ARn model is possible through this formulation using the original A and C matrices of the SKF model to build \mathcal{A} . For further details refer to [24].

3.3. Masking high amplitude movements

Sudden and large amplitudes movements can lead the eye pupil to be only partially seen by the wavefront sensor, producing poor quality measurements. The identification of such events are not only important for the post-processing (in which images acquired during that particular period are not taken into account: they are generally completely blurred), but also for the real-time parameter estimation process. Bearing Fig. 2 in mind, we have set an automatic procedure for identification of abrupt changes in the trajectories, which were masked out in real time during the RLS parameter estimation process. Hereafter we consider that the PT is working at a 80 Hz frame rate, and the masking is designed for this sampling rate, in real time. The same mask is used in the EKF estimation, where we freeze the state vector during such events; this is done to have a fair comparison basis for the methods.

Once a sudden and large amplitude movement is identified, the masking is done in both axes of the PT data (i.e. for horizontal and vertical movements). At time k , the standard deviation σ_k of all δp^m up to that moment is computed, and a flag of 'non-useful data' is ascribed to a temporal window comprising $k \pm 2$ if $|\delta p_k^m| > 3.5\sigma_k$. This procedure masks out up to about 16% of the trajectories in our data sample. Absolute displacements beyond $3.5\sigma_k$ are rare when fixing firmly a target and can be masked in the post-processing (i.e. discarding the corresponding science images; see Section 5 on wavefront errors).

4. Experimental data: pupil displacement prediction

4.1. Description of prediction error rms data

Adopting the state vector defined in Section 3, we compare here the measured and predicted positions for multiples of $\Delta t = 12.5$ ms, the sampling period of the PT. We call horizontal (H) and vertical (V) the two position axes measured by the PT, in order to avoid confusion with the variables x and y used previously. Performance of the dummy predictor and of the predictors based on SKF+RLS and on EKF was evaluated by computing, for each of the 52 trajectories, the rms (in microns) of the prediction errors for horizontal and vertical movements and for prediction horizons ranging from one to five time steps. Table 1 presents the mean rms value, standard deviation, maximum value for each direction and prediction horizon, together with the success rates for both Kalman filters, i.e. the percentage of trajectories for which they perform better than the dummy predictor.

4.2. Improvement in mean rms

For most of the 10 combinations of direction and prediction horizon, both Kalman predictors give lower average rms and standard deviation than the dummy predictor (the extended Kalman filter has higher average and standard deviation for horizontal displacements and prediction horizons of 4 and 5 steps). The Kalman filters also have lower maximum rms in all 10 cases. However, dummy and Kalman predictors values are not so different (they differ on average by less than $2 \mu\text{m}$, i.e. much less than the PT measurement error itself), and it is necessary to perform a statistical test.

For the two Kalman predictors, for every direction (horizontal and vertical) and prediction horizon (1 to 5 steps), we performed a standard one-sided t -test (based on Student's statistics, using the Matlab function `ttest.m`) to decide whether the mean rms of the prediction error for SKF+RLS or EKF is significantly lower than the mean rms for the dummy predictor. Table 2 presents the p -values for each of these 10 tests (i.e., the probability of observing a Student test statistic as extreme as, or more extreme than, the observed value under the null hypothesis that the two mean rms are equal). For vertical displacements, both Kalman filters perform significantly better (at a 5% confidence level) than the dummy predictor for all prediction horizons.

Table 1. Comparison of the performance of the predictors for all trajectories. Each cell gives the mean, standard deviation and maximum value of the series of 52 horizontal and vertical prediction errors rms (in μm) computed for all trajectories; the success rate (r) of both Kalman predictors over the dummy predictor for the 52 trajectories is displayed in the two last lines. Saccades and abrupt changes in the trajectories have been masked out in the estimation process, as described in the text. H: horizontal; V: vertical. $\Delta t = 12.5$ ms.

Predictor	1 step (Δt)		2 steps ($2\Delta t$)		3 steps ($3\Delta t$)		4 steps ($4\Delta t$)		5 steps ($5\Delta t$)	
	H	V	H	V	H	V	H	V	H	V
Dummy										
mean rms	8.57	6.39	15.50	10.40	21.28	13.90	28.58	21.11	35.58	24.72
st.d.	3.29	2.27	6.10	3.99	8.58	5.62	11.99	14.34	16.35	13.34
max.	18.81	15.15	34.90	25.89	47.82	35.49	65.07	82.54	91.51	88.46
SKF+RLS										
mean rms	6.65	6.02	13.51	9.92	19.98	13.49	27.78	20.74	35.37	24.36
st.d.	2.65	2.19	5.36	3.51	7.98	5.01	11.28	14.08	14.78	12.96
max.	14.11	13.34	30.01	23.03	43.18	32.62	59.22	81.12	84.22	87.01
EKF										
mean rms	6.76	5.95	13.81	9.84	20.54	13.41	29.01	20.76	37.25	24.44
st.d.	2.69	2.19	5.54	3.60	8.36	5.10	12.12	14.12	16.83	13.04
max.	14.33	13.72	31.67	23.59	46.37	33.29	64.56	81.93	88.48	87.83
r (SKF+RLS)	92%	69%	85%	65%	56%	60%	62%	56%	56%	54%
r (EKF)	92%	73%	83%	65%	52%	63%	52%	56%	42%	48%

For horizontal displacements, SKF-RLS and EKF perform significantly better than the dummy predictor for horizons ranging respectively from Δt to $4\Delta t$ and from Δt to $3\Delta t$.

Table 2. *p*-values for Student tests of assumption ‘the mean rms of SKF+RLS or EK prediction error is smaller than the mean rms of the dummy predictor’ for horizontal and vertical displacements ranging from Δt to $5\Delta t$. Values below 0.0001 are rounded to zero. Any value above 0.05 indicates that the difference is not significant.

Test	Δt		$2\Delta t$		$3\Delta t$		$4\Delta t$		$5\Delta t$	
	H	V	H	V	H	V	H	V	H	V
SKF+RLS vs Dummy	0	0	0	0.0018	0.001	0.0087	0.035	0.0146	0.3528	0.0149
EKF vs Dummy	0	0	0	0.0002	0.0378	0.0087	0.7584	0.0096	0.9475	0.023

Better results are obtained for horizontal displacements predictions for all methods, with smaller average rms, standard deviations and maximum values. This is in agreement with [28] where the authors noticed that, for fixational eye movements, the horizontal components are much strongly correlated than the vertical ones at the short time scale. They explain this by the fact that micro-saccades are controlled by different brainstem regions, as reported in [29].

Also, larger prediction horizons lead to non significant differences: the performance of the Kalman filters for large horizons is not improved with respect to the dummy one. This is not surprising, as the models of the Kalman filters have been built simple and for short-term prediction (at most $2\Delta t$). Examples with good rms improvement and without rms improvement are given in Fig. 5 and 6 respectively, where a zoomed part of one trajectory of the horizontal position is plotted along with predictions for Δt and $2\Delta t$. The solid black line denotes the measurements, whereas the dashed one represents the different predictions. Longer horizons introduce larger gap between predicted and measured trajectories, but also spurious peaks, especially in the Kalman cases as illustrated by Fig. 5 where a peak starts to appear between 3

and 3.2 s for Kalman predictors with $2\Delta t$, increasing significantly the rms.

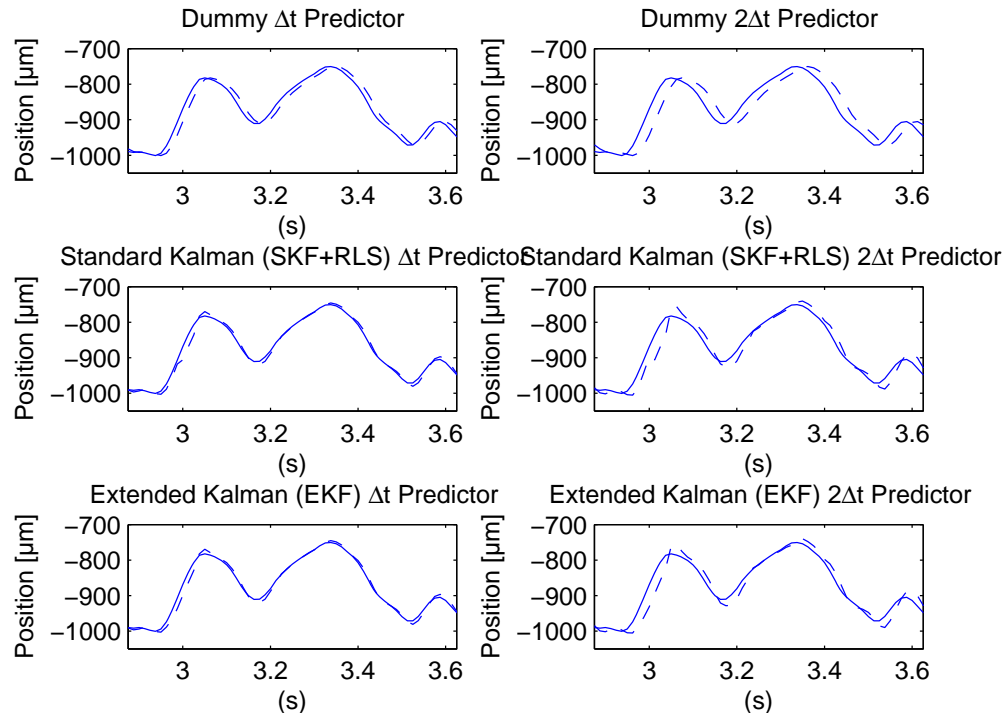


Fig. 5. Comparison of predictors for Δt and $2\Delta t$, in a good improvement case. Solid line: measurements; dashed line: predictions.

4.3. Comparison of success rates

Table 3 is another way to look at the results, by taking the success rate of the Kalman based predictions over the dummy prediction in the trajectories themselves. For each horizontal and vertical displacement trajectory, we compute the percentage of sample times where the Kalman prediction is closer to the measurement than the dummy prediction (in terms of absolute value of the prediction error for each sample). These percentages are then averaged over the 52 trajectories, and standard deviation, maximum and minimum values are given. Any value above 50% in the mean percentage indicates a prediction improvement in average over the dummy predictor. Maximum and minimum values when Kalman predictors are better than dummy lie within the ± 2 standard deviation interval, so there are no outliers. The first maximum value of 77% (SKF+RLS, H, max. when better) means that over all 52 trajectories, for the horizontal displacements, the prediction with SKF+RLS is better than dummy for at most 77% of the samples. Similarly, the first minimum value of 40% (SKF+RLS, H, min. when better) indicates that the prediction with SKF+RLS is better than dummy for at least 40% of the samples.

4.4. Relative performance improvements and degradations

Finally, it is interesting to appreciate one- and two-step prediction performance in terms of rms when Kalman predictors perform better than dummy (for SKF+RLS H and V, 92% and 69% of the cases, see the last two lines of Table 1 for all the percentages), and when dummy performs better than Kalman predictors (for SKF+RLS H and V, 8% and 31% of the cases). We thus

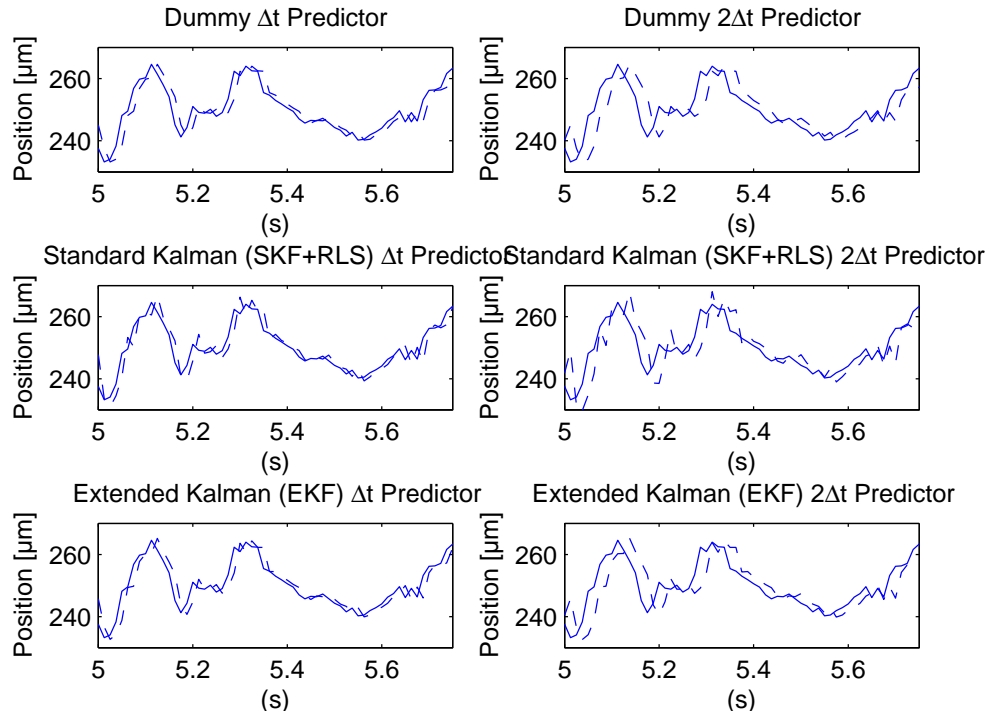


Fig. 6. Comparison of predictors for Δt and $2\Delta t$, in a case without improvement. Solid line: measurements; dashed line: predictions.

Table 3. Prediction error along trajectories: mean, standard deviation, maximum and minimum values of the percentage of sample times in which the Kalman filters give a smaller prediction error than the dummy predictor. Any value above 50% of mean percentage indicates a better prediction than with dummy.

	Δt		$2\Delta t$		$3\Delta t$		$4\Delta t$		$4\Delta t$	
	H	V	H	V	H	V	H	V	H	V
SKF+RLS vs Dummy										
mean percentage of samples	65%	53%	59%	53%	54%	50%	50%	49%	48%	49%
stand. dev.	7.7%	8.7%	8.3%	7.7%	9.1%	7.3%	9.2%	6.9%	9.3%	6.8%
max. when better	77%	72%	72%	74%	67%	74%	65%	74%	63%	74%
min. when better	40%	37%	41%	37%	37%	35%	34%	38%	30%	35%
EKF vs Dummy										
mean percentage of samples	67%	57%	61%	56%	56%	55%	53%	54%	50%	53%
stand. dev.	6.8%	6.8%	7.6%	6.3%	8.5%	6.2%	8.8%	6.2%	8.9%	6.1%
max. when better	79%	71%	74%	75%	71%	76%	69%	76%	67%	76%
min. when better	43%	41%	46%	43%	43%	44%	37%	44%	34%	43%

compare, in Table 4, (i) the relative performance improvement obtained by the Kalman predictors when they perform better than dummy, and the relative performance degradation when they perform worse, and (ii) the relative performance improvement obtained by the dummy predictor when it performs better than Kalman predictors, and the relative performance degradation when it performs worse. Maximum values lie within the mean ± 3 standard deviation interval (except for the first value of 95% in the last line).

The values presented here confirm that in average, Kalman predictors perform better than

dummy: for one-step ahead prediction, the mean improvement for 92% of the horizontal trajectories is around 24%, and is around 11% for about 70% of the vertical trajectories. The performance degradation for these trajectories when dummy is worse than Kalman predictors is about 33% and 14% for horizontal and vertical trajectories respectively, while maximum degradation is very high (95% and around 44% respectively). For the trajectories where dummy has better rms (8% for horizontal and around 30% for vertical, as deduced from Table 1), the improvement is very low (at most 7%), while performance degradation on these trajectories induced by Kalman predictors is very low.

Table 4. Relative performance improvements and degradations: mean, standard deviation and maximum value of rms over the 52 trajectories in percentage points and μm (in brackets). For dummy with respect to (w.r.t.) SKF+RLS for example, $p\%$ of improvement means that dummy decreases the prediction error rms of SKF+RLS by $p\%$, and $p\%$ of degradation means that dummy increases the prediction error rms obtained by SKF+RLS by $p\%$. Horizontal and vertical displacement are considered, with one- and two-step ahead predictions.

	Δt		$2\Delta t$	
	H	V	H	V
Improvement with SKF+RLS w.r.t. dummy				
mean	24% (2.084)	12% (0.753)	15% (2.849)	8% (0.986)
standard deviation	12% (1.257)	9% (0.719)	9% (1.658)	7% (1.116)
maximum	49% (5.294)	30% (3.132)	40% (6.494)	26% (4.615)
Degradation with SKF+RLS w.r.t. dummy				
mean	2% (0.076)	7% (0.472)	4% (0.416)	5% (0.455)
standard deviation	2% (0.074)	3% (0.301)	3% (0.352)	3% (0.314)
maximum	4% (0.181)	12% (1.129)	8% (1.955)	11% (1.11)
Improvement with EKF w.r.t. dummy				
mean	23% (1.985)	10% (0.663)	14% (2.242)	8% (0.949)
standard deviation	12% (1.251)	9% (0.693)	9% (1.590)	8% (1.112)
maximum	49% (5.321)	32% (3.203)	34% (6.547)	29% (4.938)
Degradation with EKF w.r.t. dummy				
mean	4% (0.294)	2% (0.159)	8% (0.962)	2% (0.163)
standard deviation	3% (0.242)	3% (0.313)	11% (1.488)	3% (0.354)
maximum	7% (0.599)	13% (1.221)	30% (4.491)	13% (1.561)
Improvement with dummy w.r.t. SKF+RLS				
mean	2% (0.076)	7% (0.472)	4% (0.416)	5% (0.455)
standard deviation	2% (0.074)	3% (0.301)	2% (0.352)	3% (0.314)
maximum	4% (0.181)	10% (1.119)	8% (1.195)	10% (1.111)
Degradation with dummy w.r.t. SKF+RLS				
mean	34% (2.084)	14% (0.753)	19% (2.429)	10% (0.986)
standard deviation	21% (1.257)	13% (0.719)	14% (1.658)	10% (1.116)
maximum	95% (5.294)	43% (3.132)	68% (6.494)	35% (4.615)
Improvement with dummy w.r.t. EKF				
mean	4% (0.294)	2% (0.159)	7% (0.962)	1% (0.163)
standard deviation	3% (0.242)	3% (0.313)	9% (1.487)	2% (0.354)
maximum	6% (0.599)	11% (1.221)	23% (4.491)	11% (1.561)
Degradation with dummy w.r.t. EKF				
mean	32% (1.985)	13% (0.663)	18% (2.242)	10% (0.949)
standard deviation	20% (1.252)	13% (0.693)	13% (1.590)	10% (1.112)
maximum	95% (5.321)	46% (3.203)	50% (6.547)	42% (4.939)

4.5. Pupil displacement prediction: summary of results

Our experimental results confirm that a better prediction of the pupil displacements can be achieved with the use of the Kalman tools. This improvement (measured as the mean rms of

the prediction error) is statistically significant for all vertical displacements and for horizontal displacements of less than $4\Delta t$ (with a slight advantage to the SKF+RLS method). However, this also suggests that even at a 80 Hz PT rate, the measurements are weakly correlated, so the Δp process at this rate tends to be close to a random walk, leading to results only slightly better than the ones obtained with the dummy model. The goal of the next section is to evaluate the impact of pupil motion and of course the level of improvement brought by the predictors presented above.

5. Experimental data: impact of pupil motion on residual wavefront errors

In order to evaluate the impact of pupil motion on the imaging quality, we have performed a series of simulations of the displacement of eye aberrations. The goal is to evaluate, given a frozen aberration, the contribution to the residual wavefront due only to pupil displacement (moving aberration hypothesis). Notice that, as we are focused on residual wavefront budget error, the hardware limitations are not taken into account and the frozen aberration is assumed to be known (i.e. there is no fitting error for the correction and aberrations are perfectly represented — no spatial discretization due to a wavefront sensor).

5.1. Generation of phase screens

Phase screens were generated using sets of normalized Zernike coefficients obtained from wavefronts measured on real healthy subjects by Imagine Eyes. The coefficients ranged from radial orders 2 to 6, totaling 25 modes.

The phase screens were generated inside a larger frame than the actual pupil size of the instrument, so we could always see aberrations when shifting the phase screens to mimic movements. When setting up the size of the phase screen frame, the statistics of the eye movements (Section 2) were taken into account, and extreme eye movements were actually not considered in the final computation of the rms (as the images acquired during such events can be discarded in post-processing).

The statistics of the generated phase screens were compared with a parameterization given in the literature. Based on Fig. 9A of [22] and a rough estimate of its parameters for our phase screen size, let us assume that the aberrations have a dependence of wavefront variance σ_n^2 (this is a ‘partial’ variance in the sense that is due only to a radial order mode) vs. Zernike radial order (n) as $\sigma_n^2 = 8.96 \exp(-1.5n)$. Using this expression, synthetic phase screens can be generated by taking a set of normalized Zernike modes (Z_k) and linearly combining them with Gaussian random weights with zero mean and $[\sigma_n^2/(n+1)]^{1/2}$ standard deviations. The histograms of synthetic and real data wavefront rms are shown in Fig. 7. The distributions are not consistent (two-sample Kolmogorov-Smirnov test (Matlab function `kstest.m`) at 5% significance level) due to the large scattering of the real data, but have consistent median values (as attested by the Mann-Whitney U-test (Matlab function `signtest.m`) at 5%). We have considered that this data set obtained from measured wavefronts was more reliable for performance evaluation than synthetic data, and have used it for our performance analysis.

5.2. Wavefront error analysis

The advantage of a PT as the one analyzed here is that it can run at a faster rate than the wavefront sensor, as mentioned in [16]. Any analysis of error budget here requires therefore the computation of differences between a reference phase screen (the corrected one) and a displaced phase screen. This difference has to be piston and tip-tilt corrected, because these modes introduce only an image translation. Since the final science image is produced by stacking individual frames, the translation can be easily accounted for in the post-processing. In this section we focus on the impact of the PT frame rate and of the prediction models on the error budget.

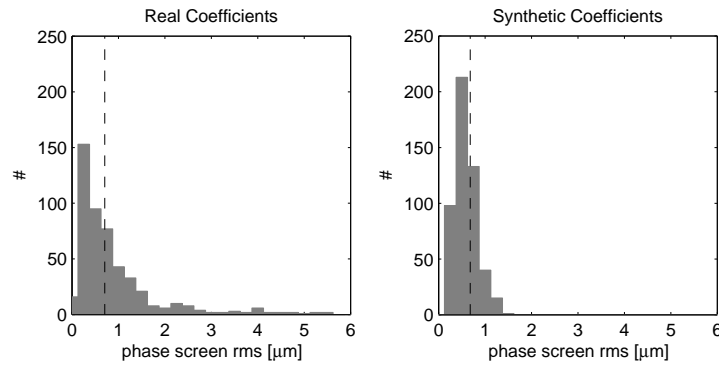


Fig. 7. Statistics of phase screens generated from real and synthetic Zernike coefficients. The latter shows a more concentrated distribution, while the histogram generated from real data exhibits a large scattering. Dashed lines denote median values.

5.2.1. Impact of PT frame rate on wavefront error budget

Once the phase screens have been generated, it is possible to predict the wavefront error resulting from a random pupil displacement. Figure 8 was built based on 50 phase screens randomly selected amongst 500; for each of them, 10 random directions were generated for every displacement amplitude; displacement amplitudes ranged from 0 to 10% of the pupil size, by steps of 1%. The reference phase screen is kept centered, and the rms of the difference between this latter and the displaced pupil is computed, after piston and tip-tilt modes have been removed. The figure shows average rms values along with their standard deviations (curve and error bars, respectively). The trend of increasing error follows a quite linear relation, and shows that for a 10% of pupil size displacement not accounted for by the system correction, the wavefront error can reach, on average, about 140 nm, with a large scatter around this value (bars correspond to $\pm\sigma$, with σ the empirical standard deviation computed over 500 values).

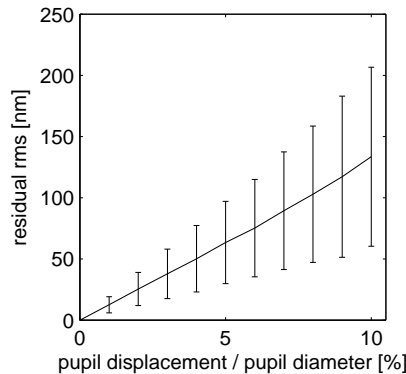


Fig. 8. Impact of random displacements of the pupil on the residual rms. The difference between reference and shifted phase screens are piston and tip-tilt subtracted before the computation of the rms values. The curve represents the average value and the error bars are the standard deviations calculated from 500 realizations (a number of 50 randomly selected phase screens were used here, and 10 random directions used for each displacement amplitude).

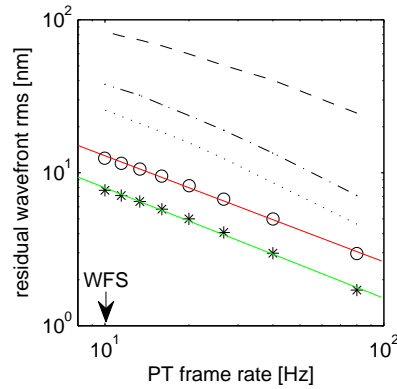


Fig. 9. Residual wavefront error (in rms, piston/tip/tilt removed) as a function of the PT frame rate. The curves correspond to upper limits considering all 99%, 95% and 90% smallest absolute value of the one-step displacements (dashed, dot-dashed and dotted lines). The mean and median computed over all one-step displacements are shown as circles and stars, respectively. The red and green lines represent the linear fits as explained in the text. The arrow shows the current WFS rate.

We can connect this result with the displacements statistics of Section 2, Fig. 4. Those statistics were built based on the delay between PT position measurements, so such delays can be considered as the inverse of the PT frame rate. The expected residual wavefront rms as a function of this rate is shown in Fig. 9, which has been obtained by combining Fig. 8 and 4. For the latter, both horizontal and vertical displacements were considered together (calculating the square root of the sums of the squared displacements in both directions), instead of using the statistics for separate axes. The sampling frequency of the WFS is shown in the figure as a reference point. It corresponds in our set-up to the case without PT, that is, a delay of $5\Delta t$ between WFS measurement and DM correction (and thus image acquisition), which gives an eye pupil position sampling frequency of 16 Hz. We used the mean/median rms values and made a linear fit on the logarithmic scale. From that we get a general expression for the residual rms expected for a given PT frame rate (in Hz):

$$\text{wavefront rms} \approx \alpha \times (\text{PT frame rate})^\beta [\text{nm}] \quad (28)$$

with $\alpha = 63.8$ and $\beta = -0.69$ using the mean wavefront rms, and $\alpha = 42.1$ and $\beta = -0.72$ using the median ones. Such fits are shown in red and green lines in Fig. 9, respectively. The parameters have been estimated from the 500 realizations mentioned in Fig. 8, and can vary a bit around such values depending on the phase screens randomly selected for the simulations (about ± 3 for α and ± 0.01 for β).

From this analysis, we confirm that, in order to minimize the residual error, we have to keep the PT frame rate as high as possible. The power around -0.7 obtained above tells us that we double the mean rms error, if for example, we drop the PT frequency from 80 Hz to 30 Hz. The curves in Fig. 9 thus chart the gain in performance to be expected from the use of a PT at a faster rate than the WFS. The case where the PT rate is equal to the WFS rate (about 10 Hz) indicates the level of performance that could be achieved without PT. It shows that for 95% of the one-step displacements, the performance loss in terms of wavefront rms is lower than 50 nm, (and lower than 100 nm for 99% of the displacements). This is to be compared with the rms of the wavefront itself (without displacement), which lies between 0.1 and 8 μm approximately. Depending on the trajectory, the rms of the aberration caused by the eye displacement may thus

represent a large percentage of the total aberration.

In [16], the authors present an AO system running at about 10 Hz (8.4 precisely), and the AO loop led to an average rms of $0.12 \pm 0.05 \mu\text{m}$ for the three studied subjects (a PT was used in this experiment at the same sampling frequency of 8.4 Hz to verify that eye aberrations were mainly frozen wavefronts moving with the eye). If we suppose that the error budget due to the eye displacement is at this frequency at most 40 nm rms (the value at 95% in Fig. 9), this represents the third of $0.12 \mu\text{m}$. With a PT running at, e.g., 80 Hz, this budget could decrease to 7 nm rms (the 95% value at 80 Hz), which lies in the error range of the experiment. For an imaging wavelength of 850 nm, going from 120 nm rms to $\sim 113 \text{ nm}$ ($= \sqrt{120^2 - 40^2 + 7^2}$) would lead to an increase of about 5 points of Strehl ratio (SR), which is very significant in terms of image quality. (The Strehl Ratio was computed with the Marechal approximation $\text{SR} \approx \exp(-\sigma_{\phi_{\text{rms}}}^2)$, where the variance in radians is given by $\sigma_{\phi_{\text{rms}}}^2 = (\frac{2\pi}{\lambda_{\text{im}}} \text{rms})^2$ with λ_{im} the imaging wavelength.)

5.2.2. Impact of position predictions on wavefront error budget

With a PT working at 80 Hz, we can check now the impact of the position predictions on the wavefront errors. These can be obtained by computing the root mean square of the difference between a reference wavefront displaced to the *real* position p and the same wavefront displaced to the estimated position \hat{p} , at each PT frame. This reflects the wavefront error contribution due to the position estimation error, if computed inside the fixed and centralized pupil of the imaging camera, after piston and tip-tilt removal. The larger phase screen frames have been generated with a resolution of about $6 \mu\text{m}/\text{pixel}$, so they were well-suited to track the impact of small prediction errors in the positions. The simulations consist in computing the average rms of displaced screens difference along a whole trajectory, and this is done for 20 phase screens. Absolute displacements of the phase screen larger than about $500 \mu\text{m}$ have not been taken into account in the computation of the rms error, as the corresponding images would be discarded. The histograms of the overall rms average values are shown for the 52 pupil trajectories in Fig. 10 for prediction horizon Δt .

Mean values of the wavefront errors rms shown in the histograms as dashed lines are gathered in Table 5. Comparing such values with the non-PT case in Fig. 9, we see once again that the performance improvement is important when using a PT with fast rate. In the case of a fast rate, predicting the pupil position can be made simple by keeping the last measured value. Using more complicated schemes does not bring here significant improvement (less than 0.5 point of SR for the numerical example presented in the previous section).

Table 5. Mean wavefront residuals (nm rms) obtained from the prediction models, for three different time delays. These numbers are to be compared with the non PT case, corresponding to an approximate delay of $5\Delta t$ and which leads to an mean and median rms of 12.1 nm and 7.7 nm, respectively.

	mean rms			median rms		
	Δt	$2\Delta t$	$3\Delta t$	Δt	$2\Delta t$	$3\Delta t$
Dummy	2.11	3.66	5.01	1.91	3.33	4.51
SKF+RLS	1.70	3.18	4.80	1.59	3.12	4.54
EKF	1.73	3.22	4.70	1.58	3.02	4.37

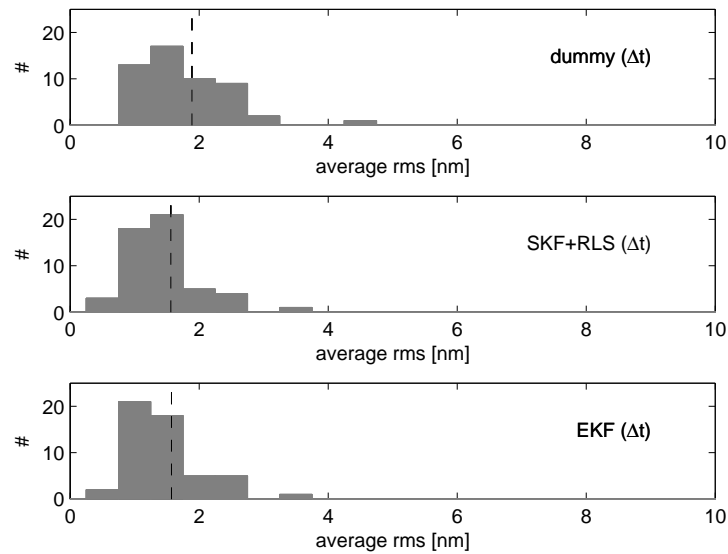


Fig. 10. Residual wavefront errors obtained from prediction models, expressed in rms (nm), for a prediction delay of Δt . The dashed line denotes the median value of the histogram.

6. Discussion and conclusions

We have investigated in this work the impact that a pupil tracker working at different rates than a WFS can bring to a retinal imaging instrument equipped with adaptive optics. Using a data set of 52 trajectories provided by Imagine Eyes, we started by analyzing the statistics of the eye movements, from which we could derive constraints both for real-time and post-processing of science images. Through this analysis, it has been verified that the vast majority of pupil one-step displacements lies (in absolute value) below $200\ \mu\text{m}$, and on average below $30\ \mu\text{m}$ for measurement delays inside the AO cycle.

Three models have been presented for pupil position estimation, assuming a pupil tracker working at 80 Hz. The first was the ‘dummy’ one, which corresponds to keeping the latest measurement as a predictor for the next position (and which represents the best solution for the case of a random-walk nature for the one-step displacements). The two other predictors are of Kalman filter type, based on an auto-regressive modeling (order 1) of the position, with parameters estimated in real time either through a recursive least square (RLS) approach or by including the parameter in the state vector (extended Kalman filter, EKF, approach). Extreme displacements have been masked out automatically in the real-time process, since they represent the abrupt changes known as saccades.

Although one-step pupil displacements trajectories are weakly correlated, the pupil position prediction obtained using identification/prediction tools such as Kalman filters can improve results with respect to the dummy solution in terms of position prediction error rms. Improving further prediction performance would need to consider more complex models and predictors, capable of detecting and adapting to the different displacements behaviors in real time.

We have used 500 phase screens from healthy subjects, obtained by linearly combining Zernike modes with measured aberration coefficients provided by Imagine Eyes. The idea was to test the impact of the position predictors on the wavefront error, neglecting any other external contributors to the system error budget. The difference between a reference phase screen and the displaced one gives, inside the system’s pupil, the basis for the calculation of the wavefront

error.

Considering a fixed reference at the pupil center, we have checked how pupil displacement impacts on the wavefront error. The statistics of the absolute value of the one-step displacements help us to check the impact of the pupil tracker frame rate on the wavefront error. We have derived a relation for the expected wavefront error rms with pupil tracker rate proportional to the power of -0.7 . This means, for instance, that if the pupil tracker rate drops from 80 Hz to 30 Hz, one can expect the rms wavefront error to double. Comparing what is expected from a pupil tracking at 10 Hz (which corresponds to an ideal AO set-up at 10 Hz with perfect WFS measurement and DM correction) with a pupil tracking at 80 Hz, the performance improvement in terms of wavefront rms is significant. We have compared the error budget caused by the pupil displacement with real experiments presented in [16], for a system running at around 10 Hz. It is shown that for such a system, the maximum degradation impact due to pupil displacement when compared with the use of a PT at 80 Hz may be significant. The numerical example led to an error budget of a third of the given experimental wavefront rms, and a Strehl ratio decrease of about 5 points, significant in terms of image quality.

Rates higher than 80 Hz have not been considered, as the one-step displacements have at this rate a value in the range of the PT measurement noise (20 μm). Other PT data with higher rates, as in [30] where aberration dynamics are studied at 236 Hz, and lower measurement noise could allow to extend the results to higher PT sampling frequencies.

We have also checked the impact of the pupil position predictors on the residual wavefront, assuming a pupil tracker rate at 80 Hz. Comparing performance of the one- and two-step Kalman predictors with respect to the dummy one, we found no significant improvement in average in terms of residual wavefront rms. In this sense, just keeping the latest measurement should be enough for a good wavefront prediction. However, for other new imaging systems like AO SLO/OCT, the methods presented here could be of interest to compensate for motion in-frame distortions for example, as reported in [31].

Further developments include the establishment of the error budget for a complete system. The error budget due to pupil movement needs then to be completed with all the other budget terms (WFS measurement errors, deformable mirror fitting error, error due to WFS spatial discretization, precision of the fixed aberration estimation, calibration errors). Only residual wavefront and the fixed ocular aberration estimation depend on pupil movement. Taking into account all the remaining sources of rms wavefront error is essential to assess a complete performance for a particular AO system. The next step is then to perform experimentations combining AO retinal imaging with a PT device operating at a higher rate, and to compare retinal images in term of visual quality. This is clearly beyond the scope of this paper and left for future research. The results presented here allow nevertheless to evaluate, for any system, the possible improvement brought by a pupil tracker at different rates, when the static wavefront is supposed to be known (under the moving aberration assumption as proposed in [16]). In that sense, they are independent of system components and estimation algorithms all together, and can thus be used as part of a global system error budget analysis.

Acknowledgments

This work has been supported by French National Research Agency (ANR) through contracts ANR-09-TECS-009-03 iPhot and ANR-12-TECS-0015 ReVeal. The authors thank Betul Sahin – who developed the non-commercial PT device with Imagine Eyes® (Orsay, France) – for the pupil tracking data acquisition, Imagine Eyes® for the measured wavefront and PT data, and Serge Meimon and Jean-Marc Conan (ONERA, France) for fruitful discussions.